

1-24-2012

# Designing a Better Weather Display

Colin Ware

*University of New Hampshire, Durham, [colin.ware@unh.edu](mailto:colin.ware@unh.edu)*

Matthew D. Plumlee

*University of New Hampshire, Durham, [Matthew.Plumlee@unh.edu](mailto:Matthew.Plumlee@unh.edu)*

Follow this and additional works at: <https://scholars.unh.edu/ccom>



Part of the [Computer Sciences Commons](#), and the [Oceanography and Atmospheric Sciences and Meteorology Commons](#)

---

## Recommended Citation

Colin Ware and Matthew Plumlee "Designing a better weather display", Proc. SPIE 8294, Visualization and Data Analysis 2012, 829409 (January 24, 2012); doi:10.1117/12.906213; <http://dx.doi.org/10.1117/12.906213>

This Conference Proceeding is brought to you for free and open access by the Center for Coastal and Ocean Mapping at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Center for Coastal and Ocean Mapping by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact [nicole.hentz@unh.edu](mailto:nicole.hentz@unh.edu).

# Designing a better weather display

Colin Ware<sup>a</sup> and Matthew Plumlee<sup>b</sup>

University of New Hampshire, Durham, NH, 03824

<sup>a</sup>Center for Coastal and Ocean Mapping; cware@ccom.unh.edu

<sup>b</sup>Computer Science Dept.; mdp@cs.unh.edu

## ABSTRACT

The variables most commonly displayed on weather maps are atmospheric pressure, wind speed and direction, and surface temperature. But they are usually shown separately, not together on a single map. As a design exercise, we set the goal of finding out if it is possible to show all three variables (two 2D scalar fields and a 2D vector field) simultaneously such that values can be accurately read using keys for all variables, a reasonable level of detail is shown, and important meteorological features stand out clearly. Our solution involves employing three perceptual “channels”, a color channel, a texture channel, and a motion channel in order to perceptually separate the variables and make them independently readable. We conducted an experiment to evaluate our new design both against a conventional solution, and against a glyph-based solution. The evaluation tested the abilities of novice subjects both to read values using a key, and to see meteorological patterns in the data. Our new scheme was superior especially in the representation of wind patterns using the motion channel, and it also performed well enough in the representation of pressure using the texture channel to suggest it as a viable design alternative.

**Keywords:** flow visualization, meteorology, human perception, visual channels.

## 1. INTRODUCTION

Weather maps are possibly the most common data visualizations on earth. A significant proportion of the human population views a weather map every day, either on the pages of a newspaper, on television, or a web-site. Non meteorologists may mostly be concerned with the forecast temperature and whether or not it is likely to be overcast or clear or raining. People with meteorological knowledge are interested in the atmospheric pressure, because this helps them reason about the likelihood of storms, winds or persistent sunny weather. Some people have specialized needs for information; sailors and pilots are interested in the wind speed and direction. Across all the various constituencies probably the most common variables of interest are pressure, temperature as well as wind speed and direction.

In this paper we report a project that had two goals. The first was abstract: to see if it was possible to simultaneously visualize two 2D scalar fields and a 2D vector field such that values could be read at any data point with reasonable accuracy. The problem of visualizing continuous multivariate maps is common across many disciplines including geology, physics, microbiology as well as meteorology and oceanography. A number of visualization methods have been invented to address it<sup>5,7,8,16</sup>. Some of these displays claim to represent as many as 12 variables. But there have been no evaluations of the extent to which each of the variables are actually legible in the sense of being accurately readable. A study by Hagh-Shenas et al.<sup>4</sup> did examine the ability of subjects to read values from a map using a mosaic of small color patches displaying six variables, but the data were not continuous, instead consisting of relatively large patches representing states in the continental USA.

For our present design exercise and study, we chose to concentrate on the display of meteorological data, partly because there are de-facto standards for displaying such data. High resolution meteorological maps typically show atmospheric pressure using contours at 4 millibar intervals, and the range of pressures is commonly 50 millibars over the continental USA and adjoining oceans, meaning that 12 levels of contours must be shown (although in exceptional circumstances such as hurricanes the range can be considerably greater). Temperature is commonly shown using color codes at 5 deg (F) intervals and the range can easily be 60 degrees meaning that 12 or more distinct values must be shown. Wind speed is usually shown by means of a standardized glyph called a wind barb (Figure 1). This allows 5 knot intervals to be shown. Wind barbs are capable of showing winds of Hurricane speed in excess of one hundred and fifty knots, but ranges up to 60 knots are more common meaning that at least a 12 step scale is needed. Wind direction is given by the orientation of the barbs.

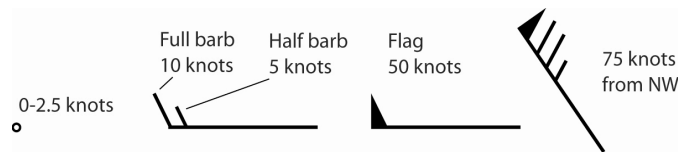


Figure 1. Wind barbs, a quantitative glyph used in meteorological maps to show weather information.

The second goal of the project was concrete: to create an example of a multivariate weather map display where each variable was provably legible to at least the resolutions given above. In addition we intended that patterns important to meteorologists be visible in the display. For example, cyclonic and anti cyclonic wind patterns, as well as areas of high and low pressure should be clearly visible. A design goal was to see if it was possible to create an effective display using different visual “channels” for the different map components according to principles derived from the theory of human perception.

From the start we planned to evaluate our display against design alternatives. We chose two. The first is based on the existing common practices of meteorologists, but with a careful re-design with the aim for finding out if existing methods are capable of meeting our design goals. The second is an example of a glyph display, included on the grounds that this type of solution has often been proposed for multivariate map displays<sup>5,7,8,16</sup>. A glyph is a graphical entity that through variations in its size, shape, orientation, texture and color can convey a number of variables. For example one variable can be mapped to color and another to the glyph size. When glyphs are used to represent a two-dimensional multivariate data field (such as is produced by a single layer of a weather model) the display map is densely packed with glyphs, either in a grid or using some other placement method. We chose to base our design on the work of Healy et al.<sup>6</sup>, mainly because their work targeted meteorological data, and their design could be used with relatively minor modifications.

In the following sections we first briefly describe the data visualized in all designs. We then present the two design alternatives (classic and glyph-based) before presenting our own design solutions which borrow elements from both and add some innovations.

## 2. NAM MODEL

The data set we chose for our evaluation study was the surface layer of the National Weather Service’s North American Mesoscale (NAM) forecast model covering the continental US. It uses a 614x428 grid in a Lambert conformal projection. Each forecast spans 84 hours in 3-hour increments (29 time-steps). A new, updated forecast is posted every 6 hours. The forecast used for the images and experiments in this paper started at 2pm EDT August 4, 2010, and ended at 2pm on August 8, 2010.

## 3. DESIGN 1: THE CLASSIC DESIGN

Standard practice in meteorological displays is to use contours for atmospheric pressure. These are often supplemented with symbols to show the centers of low and high pressure areas. The leftmost image in Figure 2 shows an example of a meteorological map showing contours for pressure, with wind barbs for wind speed and direction. When only temperature is shown on a meteorological map the most common representation is using a pseudo-color sequence, together with a key that can be read to reveal actual numbers (bottom-right image in Figure 2). The color sequence used is often based on an approximation to the physical light spectrum with purples and blues showing low temperatures, reds showing high temperatures and yellows and greens between.

Figure 3a illustrates a small sample from our *Classic* design showing the weather over northeastern Colorado. Pressure is represented using 2 millibar (mb) contours labeled at even-numbered intervals. The contours are generated using a fragment shader program that estimates distance to the nearest boundary given the underlying data. Contour labeling is accomplished by first assembling a collection of points representing the highest and lowest pressures (local maxima and minima\*) and a point for each corner of the screen. Second, the imaginary lines are run between every point in the

\* Local extrema were found for a given time step by first generating a global ranking of pressure values, and then progressing through the first 25 items in that ranking. Any that were too similar (within 2.2 mb) or too close to more extreme values were discarded.

collection, and the intersections of these lines with contour boundaries are recorded. Finally a coarse grid is used to filter out extraneous intersections: only one intersection at each contour level is allowed to remain, and the one closest to the center of the grid cell is the one that would represent that contour for the cell. The surviving intersections resulted in labels. The result does not label every contour, nor does it avoid the occasional occurrence of a label overlapping multiple contours in high-gradient areas.

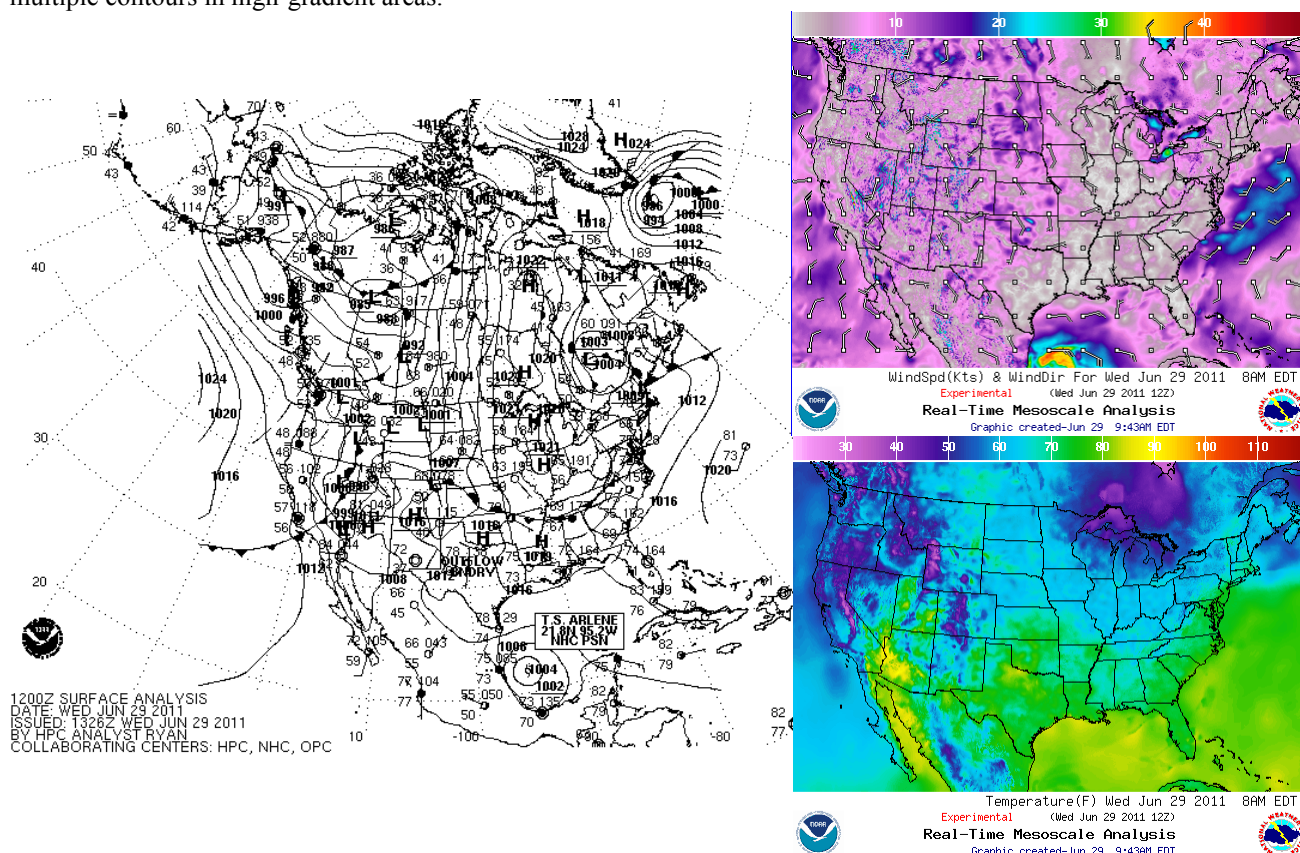


Figure 2. Images from the NOAA National Weather Service site. In the image at left, black contours show pressure contours, temperature and dew point are given with numbers, and wind speed and direction are shown using wind barbs. In the top-right image, both wind barbs and color are used to show wind speed, while the barbs show direction. In the bottom-right image, color is used to show temperature.

The *Classic* design uses traditional wind barbs to represent wind direction and speed. The wind barbs are evenly spaced 10 data cells apart. Each half-length barb represents 5 miles per hour (mph) of wind speed, while each full-length barb represented 10 mph. Wind barbs are generated to represent ranges centered on integral multiples of 5 mph, with a range of plus or minus 2.5 mph. Wind speeds of 0-2.5 mph are represented by square dots. The spacing of both the wind barbs and the pressure levels was chosen to best approximate what the spacing and resolution might be if the weather maps such as those shown in Figure 2 were designed take up a full screen.

### 3.1 The color sequence design

There has been much debate about the proper design of a color sequence for displaying a univariate map. Various researchers have argued that the spectrum approximation sequence is a poor choice on the grounds that to see the overall patterns in the data it is better to use a sequence that monotonically increases in luminance<sup>11</sup>. This is because the visual system mostly uses luminance variation to determine shape information, color providing a secondary attribute. Nevertheless, there are reasons for using a sequence that is somewhat like the physical spectrum to display temperature. Some colors have semantically intuitive mappings such as blue for cold and red for hot. But in order to get clearly distinguishable intermediate temperatures, other colors are needed and a transition through green and yellow provides many distinct steps. Also, many people are familiar with spectrum sequences on weather maps. It is capable of accurately conveying more values than any straight line sequence through color space<sup>11</sup>.

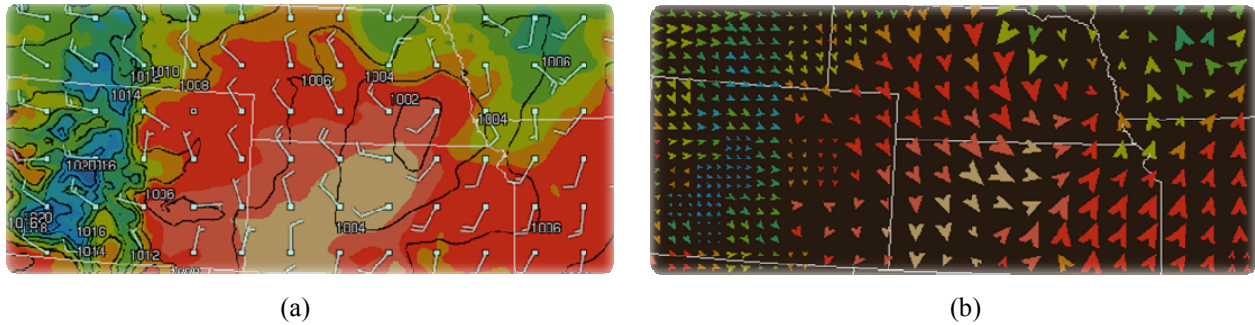


Figure 3. (a) The *Classic* design. Labeled contours are used for atmospheric pressure. Wind barbs show wind speed and direction. (b) The *Glyph* design. Pressure is shown by the number of arrows in a cell, wind direction by the way the arrows are pointing, and wind speed by the relative size of arrows. Both images show the same data.

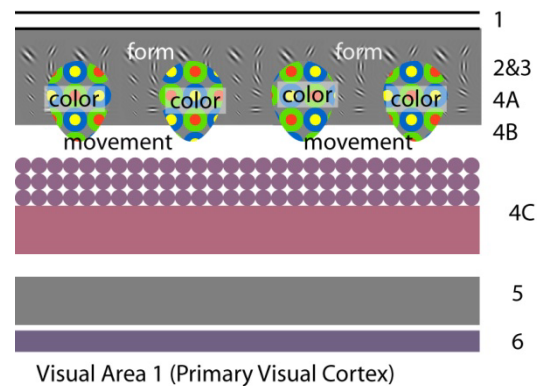


Figure 4. Adapted from Livingston & Hubel<sup>9</sup>, a cross section through the primary visual cortex (Area V1) of the brain. This shows different regions specialized to process different kinds of information in layers 2-4B: form (texture), color, and movement. The other layers deal with inputs to and outputs from this cortical region.

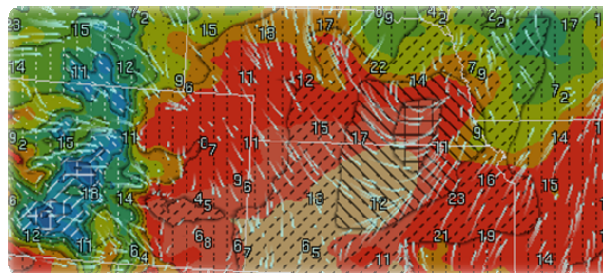


Figure 5. The *Animated* and *Static* designs. Pressure is given by contours and a sequence of textures. In the *Animated* design, streamlet animation shows the wind direction and speed. In the *Static* design, streamlets show the wind direction without animation as shown. Numbers show wind speed at sample points. This is the same data as shown in Figure 3.

Our primary requirement was to enable people to read something of the order of 12 color steps accurately, as well as to broadly show areas of high and low temperatures at a glance. The sequence we designed is illustrated in Figure 7. It has a complex path in color space; its middle section is roughly a spectrum approximation with light pink hues added to represent the hottest temperatures, and blues and grays added to represent the coldest. We do not claim that this color sequence was optimal, only that it is a reasonable solution to the design problem.

#### 4. DESIGN 2: GLYPH-BASED

The *Glyph* design is illustrated in Figure 3b. It employs regularly spaced arrowheads to represent the underlying data, in a manner suggested by Healey et al.<sup>6</sup>. The glyphs are evenly spaced 6 data cells apart, each one centered over the location of the single data point it represented—no averaging of nearby data cells is done. Temperature is mapped to glyph color using the sequence already described, and wind direction is mapped to the orientation of the delta-shaped glyphs (Healy et al. used orientation for scalar variables).

In order to represent pressure, the number of arrowheads that made up a glyph is varied using the method of<sup>6</sup> (They refer to this as “density”). For areas of pressure below 1008 mb, a single arrowhead is drawn. Between 1008 and 1016 mb, four arrowheads are drawn at half-size in each linear dimension, filling a 2x2 grid within the glyph. Finally, above 1016 mb, nine arrowheads are drawn at one-third-size in each linear dimension, filling a 3x3 grid. The center of each arrowhead remains at a fixed location, so any variation in orientation or size of the arrowheads is with respect to that fixed center.

Finally, wind speed is represented in the *Glyph* design as the relative size of the arrowheads—effectively the amount of screen real-estate that was covered by a glyph. The size of arrowheads increased linearly every 8 mph in the range of 0-40 mph, for a total of 5 different sizes. The glyph spacing was chosen to be as small as possible while still making it possible to determine wind direction at high pressures and moderate wind speeds (sacrificing low wind speeds in high pressure areas in some cases, as seen in the bottom left region of Figure 3b).

#### 5. DESIGN 3: NEW ANIMATED

The key theoretical idea behind our new design was to make use of the way the brain processes different kinds of patterns separately at early stages of processing by the visual system. Figure 4 illustrates a cross section of a small area of the primary visual cortex (V1) in the brain showing different sub-regions as classified by Livingston and Hubel<sup>10</sup>. All later stage visual processing depends on V1 where every part of a visual image is processed in parallel for a variety of different elementary features. There are three main classes of such features and these are reflected in the neural architecture: *form* (which processes local orientation and size information, providing the basis for texture discrimination), *color*, and *movement*. Sometimes these feature classes are described as visual channels, because they provide relatively independent modes via which information is transmitted to higher levels of visual processing. In addition to the neurophysiological evidence for separate processing channels there is abundant support from psychophysical studies of human perception<sup>12</sup>.

We chose to take advantage of these three basic channels to visually separate temperature, pressure, and wind variables. We chose color to represent surface temperature and in this respect our design was like both of the others. Figure 5 illustrates a static view of the new animated design.

##### 5.1 Animated wind traces

We chose motion to represent wind direction and the pattern of wind speeds because there is an intuitive mapping between moving patterns and wind patterns. To generate motion, we use a modified particle tracing method, displaying particles with fading tails representing their recent history<sup>11,15</sup>. Our use of animation should not be seen as a way of reproducing actual particle advection trajectories, but rather as a method for making wind patterns clear. We did not use a linear mapping between trace animation speed and wind speed, because this would have resulted in animation that was too fast in regions of strong winds and too slow in regions of light winds. Instead, the speed of advection is  $displayed\_speed = c \cdot model\_speed^{0.65}$ . The exponent was chosen in a pilot experiment where participants were given control over a slider to adjust the exponent parameter while the animation was running, and asked to set a value that resulted in a good representation of strong, weak, and medium wind strengths. Particles are randomly seeded over the entire field such that approximately 9600 appeared on the screen at a time. Each trace is formed by Lagrangian advection of an idealized particle according to the transformed underlying vector field. Each particle trace is given a tail that increased in transparency. To eliminate distracting visual transients at the initiation and termination of tracers, traces are randomly initiated in time. Also each individual trace begins as a point, grows to full length and terminates by fading out, as illustrated in Figure 6. The generation algorithm also results in traces that have lengths proportional to the transformed wind speed. Streamlets “die out” when they reached the edge of the screen or when they had been animated for a pre-determined (constant) lifetime. A streamlet dies by stopping advection and allowing its tail to fade completely, at which point it is “reborn” in a new random location. The initial seeding of streamlets is staggered in time so that the



dying-out and rebirth of streamlets does not happen in perceivable groupings.

In addition to animation speed and trace length as proxies for wind speed, to allow for more precise judgments we add a jittered grid of numbers showing wind speeds to two significant figures. These numbers are jittered along a grid spaced at 10 data cells (same spacing as the wind barbs in the *Classic* design. The grid is populated by first filling in wind speeds that are local maxima within a 24 data-cell area, and then filling the unpopulated grid positions with a randomly located wind speed value from within the central 4x4 region of data cells. The color of both the wind traces and the numbers is cyan, chosen so as to be easily distinguishable from any temperature coloring.

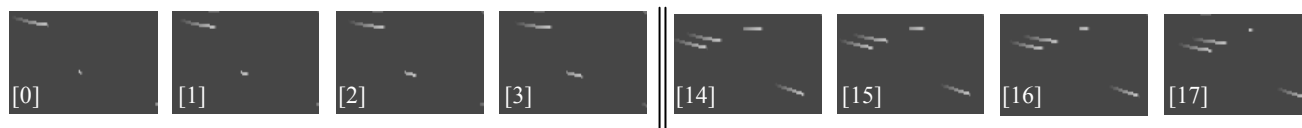


Figure 6. An illustration of the life cycle of wind traces, showing the first and last four frames of an 18-frame sequence. The first four frames show a trace coming into existence (center of each frame) and propagating forward. The last four frames show another trace (top-middle of each frame) fading out of existence.

## 5.2 Quantitative textures for pressure

We use texture to represent atmospheric pressure, designing a quantitatively coded sequence of textures<sup>13</sup>. The concept is to create a series of textures that vary monotonically in a perceptual sense from a low value to a high value and where each texture in the sequence is legibly distinct from the previous one. These textures also have the quality of laciness<sup>14</sup> meaning that other information, such as the color coded temperature, can be perceived through the interstices. This is accomplished by creating (1) a sequence of textures made of black elements where each texture has less density than the previous one, then adding (2) a sequence of textures using white elements where each elements has greater density than the previous one. The transition from black to white elements is made approximately at the average atmospheric pressure for the data shown. Texture transitions are made every two millibars. Both black and white are easily distinguishable from both the colors reserved for temperature and wind artefacts. The middle bar of the rightmost key in Figure 8 shows the entire range of textures used.

In our initial design we did not have contours. We added them because a meteorologist pointed out that pressure gradients are of critical importance in interpreting weather model output and the closeness of contours provides excellent information to support this. Therefore we added contours to provide this gradient information and to better separate the different texture bands (see Figure 5). The contours are generated similarly as in the *Classic* design, with one slight difference (as can be seen by comparing with the contours in Figure 3a). Contours in the *Classic* design are “hard” on both sides, meaning that a pixel is either black as part of a contour line, or is colored according to temperature. The contours in the *Animated* and *Static* designs are “hard” on the low side and “soft” toward the high side, meaning that the blackness of the contour tapers off through a few pixels.

## 6. DESIGN 4: STATIC VERSION OF DESIGN 3.

To evaluate the effectiveness of animation we added a variation on Design 3 in which the streamlets did not move. This was the only difference between the *Animated* and *Static* designs.

## 7. METHOD

To evaluate the four different display alternatives we carried out an experiment with two parts. The first part involved subjects repeatedly estimating the temperature, pressure, wind speed and wind direction for different points on a weather map. In the second part subjects judged how clearly they could see different features of meteorological interest.

### 7.1 Experiment Task

On each trial a cursor appeared, and the subject’s task was to estimate the temperature, pressure, wind speed and direction at that point. Figure 7 shows an example of what a subject would have seen for the *Classic* design. The subject was responsible for making selections in the key that reflected their best judgment of the underlying data at the center of a slowly spinning cursor. For example, if the cursor was centered over a boundary between light and dark green (as in

Figure 7), the subject should have selected a temperature of about 70 degrees from the key. Likewise, the subject should have traced the contour line to find the 1018 and selected 1018 mb for the pressure, and used the nearby wind barbs to estimate wind direction and speed and made appropriate selections in the key. Once all selections were done, the subject was responsible for clicking on the “Done” button in the lower-right corner.

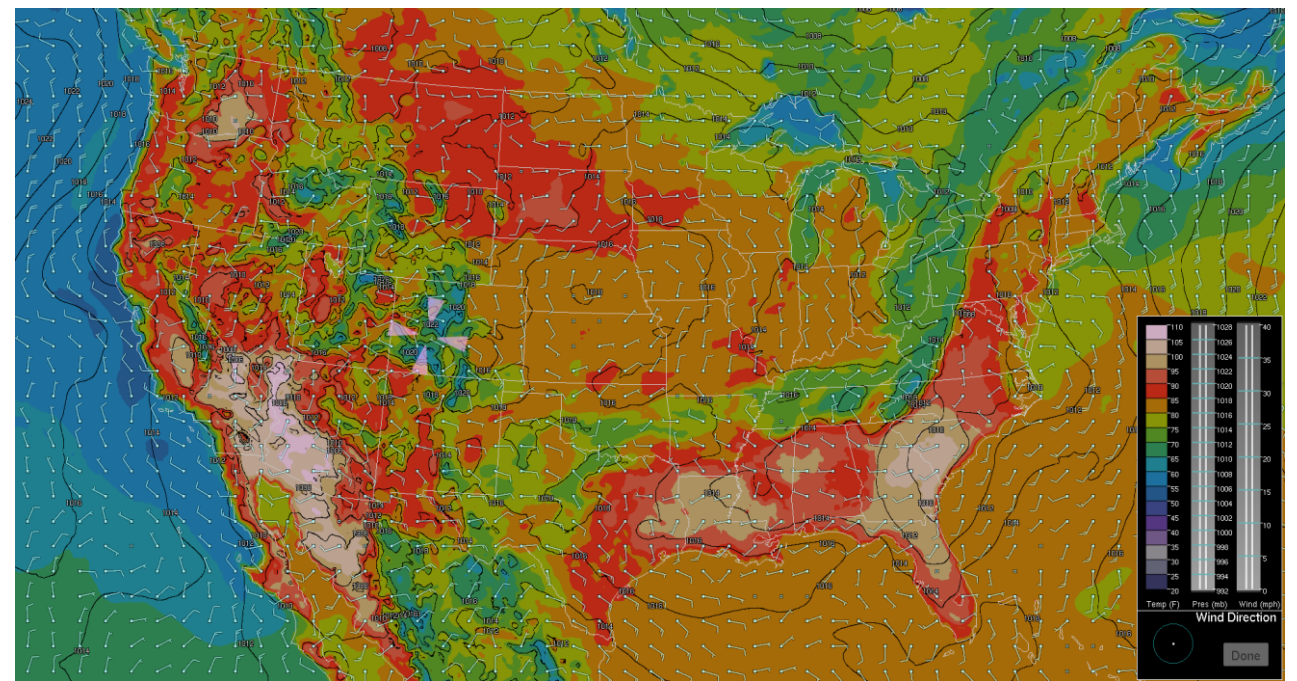


Figure 7. An example of what was seen by subjects after the one-second orientation period for the *Classic* design.

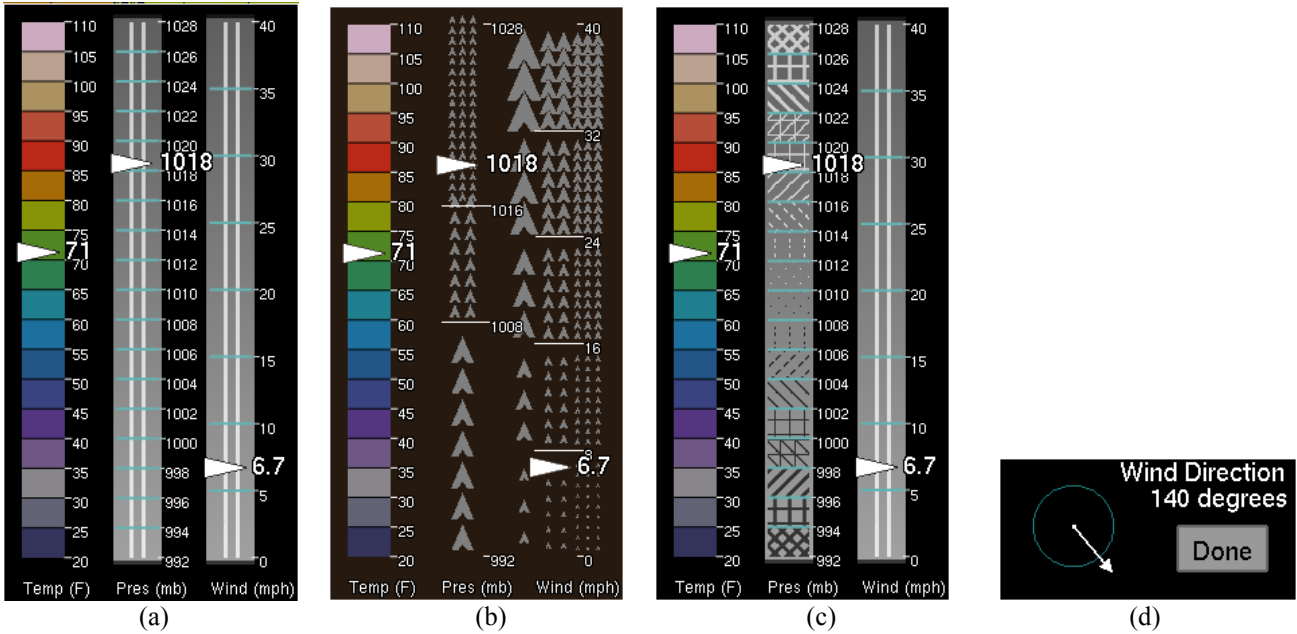


Figure 8. Keys used for the task. The upper portions of the keys presented for the (a) *Classic*, (b) *Glyph*, (c) *Animated* and *Static* designs. Each illustrates selections already made by a subject at 71°F, 1018 mb, and 6.7 mph for temperature, pressure, and wind, respectively. (d) The lower portion of the key presented for all designs. It illustrates a selection made at 140 degrees clockwise from the North position (with North pointing straight up).



The subject selected temperature, pressure, and wind speed values by using the mouse to move numeric indicators along three vertical scales in the key as in any of the keys shown in Figure 8a-c. An indicator appeared whenever the subject pressed a mouse button with the mouse cursor along a scale, and the subject could drag the indicator to the desired value. The indicator consisted of a triangle on the left and a numeric indication of the value on the right. The subject was free to adjust a selection by clicking on the scale again, at which point the indicator would “snap” to the position of the mouse cursor.

The subject selected wind direction values by using the mouse to move an arrow indicator around a circular scale, as shown in Figure 8d. The arrow indicator appeared whenever the subject pressed a mouse button near the circular scale, and it followed the mouse cursor around the scale such that an arrow always pointed from the center of the scale to the mouse. As with the vertical scales, the subject was free to adjust this selected value up until the “Done” button was pressed

## 7.2 Displays

Subjects were presented with the 4 designs during the course of the experiment. A fixed 490x271 sub-grid of the full NAM dataset was used for display (showing the continental US). The display of this data was scaled to fill the entirety of 23” widescreen LCD monitor at a resolution of 1920x1080.

Bilinear interpolation of the underlying temperature data was used to color pixels between actual data points, which tended to be roughly 4 pixels apart. For all designs, temperature was represented using an 18-color scale, as seen in Figure 8.

While the source data for the maps ran the entire gamut of the scale, subjects were most often asked to make judgments about data that was on or near land. Therefore, judgments were only made about temperatures in the 50-100° range, with rare cases in the 40°-50° and 100°-110° ranges, making the effective size of the scale roughly 10 colors. Furthermore, the random locations selected for trials never included wind speeds at or above 24 mph with this design, meaning subjects never had to make judgments using the two highest-end glyphs. Finally, the pressures that subjects were asked to make judgments on were mostly in the 1004-1022 mb range, with rare cases in the 1000-1004 and 1022-1026 ranges.

## 7.3 Experiment Design

Each experimental subject was first trained using 4 trials with each of the designs. Following this, the experiment consisted of 4 experimental blocks each containing 16 trials. Each block was partitioned into 4 sub-blocks of 4 trials where each sub-block presented a different design. Sub-blocks were arranged in a Latin squares design such that each design was first within exactly one block of trials. The Latin squares were generated randomly such that different subjects saw different orders of presentation.

For each block of trials, data for the maps presented to the subjects were selected from a different time-step of the same NAM forecast model output. At the beginning of each block, a random time-step  $t$  (in the range 0 to 28) was chosen and the first map was generated from data at time  $t$ . Each subsequent trial presented the subject with a map generated from data taken at a time 7 steps later, wrapping back around to the beginning as necessary (mod 29). Each time a block ended, the program paused to load a new time-step and the subject was given a short break until they were ready to proceed.

Each trial began by showing an outline of the United States over a black background with a key and cursor. On the first trial of each block, the following message appeared in the bottom left: “Press SPACE when you are ready to begin the next block of trials.” Once the subject pressed the space bar, the first trial of a block was like any other. The black-background map was shown for one second to make it easy for the subject to spot the magenta cursor that slowly spun in a new random location. Then this was replaced by the graphic display corresponding to the condition. The key was always present, but was updated to reflect the design for the current trial.

At the end of the one-second orientation period, the subject was presented with a fully populated map and was responsible for completing the experimental task. Task completion time was measured starting at this point, and ended when the subject clicked the “Done” button. There was no limit to the time a subject could take completing the task, but subjects were asked to try to complete the task within about 30 seconds. As a visual cue after 30 seconds, the spinning cursor would slowly pulse, gradually cycling color from red to white. The amount of time the subject took for the task was recorded along with the selections the user made in the key and the actual data values at the point indicated by the cursor.

At the end of a trial in a training block, the subject was provided with feedback on how well their answers matched with the actual values at the center of the spinning cursor. In order to minimize any training effects as well as to keep the total amount of time a subject spent on the experiment reasonable, this feedback was not given at the end of trials in the experimental blocks.

#### 7.4 Questionnaire

After all the experimental trials were completed, each subject was asked to fill out a questionnaire to rate how easily certain features could be seen. Features for the questionnaire were made in consultation with a weather expert from NOAA, but were worded so that subjects could understand. Subjects were asked to circle a dot along a 9-point scale that progressed at half-point intervals from 0 ("Can't see it") to 4 (It's very clear) for all four designs on each question. These questions asked how well the subject could see the following items:

- Q1. A large low pressure system
- Q2. A long line of converging winds
- Q3. The land-water temperature differences along a few coasts
- Q4. A moderate-sized area of (locally) highest temperature
- Q5. A small area of lowest temperature
- Q6. A moderate-sized area of (locally) highest wind speeds
- Q7. A large area of lowest wind speed
- Q8. A large-size area of (locally) high pressure
- Q9. A moderate-to-large-size area of low pressure

For each question, the subject was shown where the feature in question was (pointed to by the experiment proctor), and was given the ability to cycle through all four designs as many times as desired. Each design was labeled in the bottom-left corner of the screen, and the same labels were used on the questionnaire. For a given question, every subject saw maps of the exact same feature on the exact same time step of the same forecast as the other subjects did, but randomized elements with the *Static* and *Animated* designs made the maps appear slightly differently (locations of streamlets and jittered wind numbers).

Subjects were proctored through the experiment by a paid undergraduate research assistant.

## 8. RESULTS

### 8.1 Subjects

There were 13 subjects, eleven of whom were college students of varying backgrounds who were paid for participating. The other two were staff members of the Center for Coastal and Ocean Mapping who had little or no prior training in reading weather maps.

### 8.2 Quantitative Results

The error data is summarized in Figure 9. This shows the mean absolute errors for each of the four estimates. For temperature, the errors were 1.9 degrees or less for each of the conditions except for *Glyphs* where the errors were 2.75 degree. An ANOVA with a Tukey HSD test showed the data to fall into two groups. The *Classic* and the two new designs (*Animated* and *Static*) were all statistically indistinguishable, forming one group. The *Glyph* design was the sole member of a second group and was significantly worse than all members of the first group (all significant levels < 0.001). The greater error that occurred with the *Glyph* design is likely due to the coarser spatial discretization that resulted from the use of colored glyphs.

Errors in pressure responses showed the same pattern of results. All the methods yielded statistically equivalent results (< 1 mb mean error), except for *Glyphs* which resulted in a much greater error of around 4.7 mb. This much greater error can be attributed to Healey et al.'s design having only three levels of texture density.

The results for wind speed showed the lowest errors for the new methods (*Animated* and *Static*), while *Glyphs* were the least accurate. The mean errors for the new *Animated* and *Static* designs were 1.6 and 1.5 knots respectively; they were 2.35 knots for wind barbs (*Classic*) and 5.2 knots for the *Glyph* design. These fell statistically into three groups. The two new methods (*Animated* and *Static*) produced errors that were significantly lower than the *Classic* design using wind barbs, and wind barbs were more accurate than *Glyphs*.

The wind direction errors were the lowest for the new *Animated* wind traces (19.3 deg) with *Classic* wind barbs next (24.9 deg). Both the *Glyphs* and the new *Static* traces produced errors greater than 30.0 deg. The results did not fall into clean statistical groups according to the Tukey HSD test. Differences were significant between the *Animated* group and the (*Glyph*, *Static*) group, but they were not significantly different between the new *Animated* traces and the *Classic* wind barbs. Neither was the *Classic* design significantly worse than either member of the (*Glyph*, *Static*) group. We did note, however, that 11 out of 13 subjects were more accurate with the *Animated* traces than with the *Classic* wind barbs, and this is significant by a 2-tailed binomial test ( $p < 0.05$ ).

The average time to perform each trial was 22.2 sec. There were no significant differences depending on the display design.

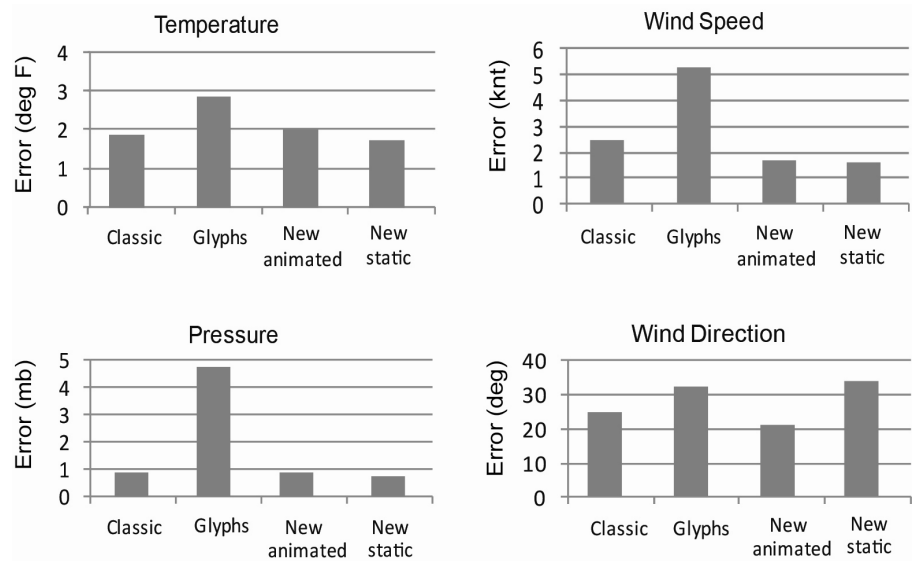


Figure 9. The mean errors for each of the measured variables for each of the four designs.

### 8.3 Qualitative Results

We now turn our attention to the subjective rating data. ANOVAs and Tukey HSD tests were run independently for each of the nine questions. Three of these yielded no significant differences. Ratings for the remaining 6 questions are given in Figure 10. In the following discussion we use the symbol “>” to mean “rated more effective than”.

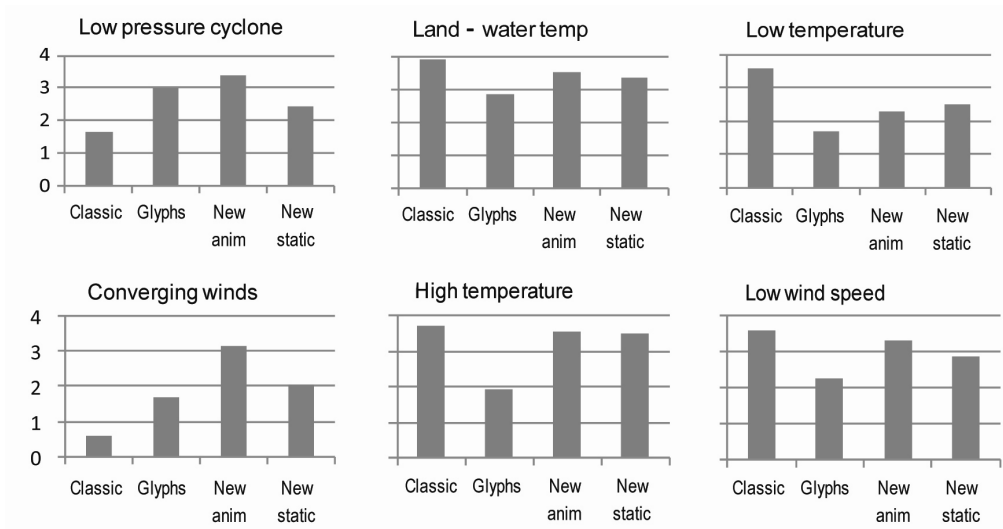


Figure 10. Subjects’ mean rating of how clearly various features could be seen with each of the four designs.

Two questions concerned wind movement patterns (Q1—*Low pressure Cyclone* and Q2—*Converging winds*). For both of them, the new *Animated* design with wind traces was rated as showing these significantly more clearly than the *Classic* design using wind barbs. The significant differences were as follows: Q1(*Animated* > *Classic*, *Glyph* > *Classic*); Q2(*Animated* > *Classic*, *Animated* > *Glyph*, *Glyph* > *Classic*, *Static* > *Classic*).

Three questions dealt with temperature patterns (Q3—*Land-water temp*, Q4—*High temperature*, and Q5—*Low temperature*). In each of these the *Classic* design was generally rated highly and the *Glyph* was rated poorly. The significant differences were as follows: Q3(*Classic* > *Glyph*, *Animated* > *Glyph*); Q4 (*Classic* > *Glyph*, *Static* > *Glyph*, *Animated* > *Glyph*); Q5(*Classic* > *Glyph*, *Classic* > *Animated*, *Classic* > *Static*, *Static* > *Glyph*).

The remaining question (Q7—*Low wind speed*) concerned the perception of an area of low wind speed. In this case there was only one significant difference: Q7(*Classic* > *Glyph*). We attribute this to the way speeds of less than 2.5 knots are displayed using wind barbs through the use of a symbol that appears very distinct when barbs appear everywhere else.

## 9. DISCUSSION

The results suggest that our new design offers significant advantages for the display of wind information over the classic use of wind barbs. Subjects judged it much better at revealing the wind front defined by a line of converging winds and also better at showing the cyclonic pattern. In addition, objective measurements revealed greater accuracy in both judged wind speed and direction. Regarding the perception of temperature and pressure, the objective results showed little difference between the new design and our re-designed version of the classic. In the case of temperature this is not surprising since in this respect the two designs used identical color sequence. The subjective measurements, however, showed that the classic was judged to show temperature more clearly than the other methods, presumably because in the new designs texture obscured more of the temperature field, while in the glyph design the temperature field was fragmented.

In the case of pressure, there were no differences between our new design and the classic, although the actual values came in one case from contours labeled with numbers, and in the other from our sequence of textures combined with a key. One way of looking at this result is that by using the sequence of textures we made numbers available for other purposes, and we chose to use them to represent wind speeds, which in turn resulted in more accurate readings of this variable. In any case, our use of texture to represent pressure proved to be a viable alternative to labeled contours with a minor tradeoff.

Overall our new design proved the best for the perception of wind patterns, but the classic did not fare as poorly as we had expected. The use of contours to show pressure and color to show temperature are clearly effective, and subjects, most of whom were unfamiliar with wind barbs, nevertheless had little trouble reading them. This suggests that weather displays like the ones illustrated in Figure 2 could be substantially improved.

The glyph-based design was significantly worse on all measurements. Would another glyph-based design have fared better? We chose the Healey et al.'s<sup>6</sup> design because it had been specifically customized to meteorological data and we could adopt it with minimal changes. But other glyph techniques can be expected to suffer from some of the same problems (for example: Kindlmann and C.F. Westin<sup>7</sup>, Laidlaw et al.<sup>8</sup>). None of these methods employed any of the techniques we used for yielding accurate measurements, such as the use of a texture sequence. Instead they all rely on continuously varying values, such as the size of glyphs. Simultaneous contrast effects apply to particle size judgments, as much as color and lightness, which means that textures where element size varies smoothly can be expected to produce large errors<sup>12</sup>.

This does not mean that glyphs should not be used. In fact our own design is substantially glyph-based although it uses a departure from the normal practice in that it uses two kinds of glyphs for different variables: the animated glyphs for wind speed and the sequence of texture glyphs used for pressure. We used additional techniques to ensure that quantities could be read: the texture sequence for pressure, the carefully designed stepped color sequence for temperature, and the numbers for wind speed.

This project began as a design exercise and it should be evident that what emerged is highly customized to the application. In so far as our design, or the classic design, is successful, this can be attributed to the specific way that data variables (temperature, pressure, wind vectors) are mapped to display attributes. As a general lesson this suggests that good designs are likely to be highly customized. It matters very much that wind is mapped to animated motion and

temperature is mapped to color, the reverse mapping would be extremely confusing. In addition much thought and many design iterations went into the exact design for the animated streamlets, for example determining that they should move at a rate proportional to the actual wind speed raised to the power of 0.65. The addition of contours to our texture sequence design came from a meteorologist's comment that showing pressure gradients is important. The designer must understand the application and the user's requirements in detail to ensure that the critical features are clear.

We feel that the use of distinct perceptual channels was successful, although of course we cannot claim to have proven this. A display is currently mounted in our lobby on a 32 inch monitor showing the *Animated* design, updating every 6 hours with the latest NAM model output. Many viewers have remarked on the effectiveness of the animated winds and how these allow for wind pattern features such as fronts to become visible. Also, we do not claim to have proven that our display is the best that could be made. But we do claim to have successfully created a map that represents two 2D scalar fields and a 2D vector field in a way that is highly legible and shows important patterns. We know of no other example that has been empirically tested to meet such rigorous requirements.

### ACKNOWLEDGMENTS

The authors wish to thank John Kelly for advice on matters pertaining to meteorology. This work was supported by NOAA Grant # NA05NOS4001153.

### REFERENCES

- [1] D. Borland, and M.R.Taylor 2nd. Rainbow color map (still) considered harmful. IEEE Computer Graphics and Applications 2007 Mar-Apr;27(2):14-7 (2007).
- [2] C. Cavina-Pratesi, R.W. Kentridge, C.A. Heywood and A.D. Milner. Separate Channels for Processing Form, Texture, and Color: Evidence from fMRI Adaptation and Visual Object Agnosia, Cerebral Cortex, 20 (10): 2319-2332. (2010)
- [3] D. Fowler and C. Ware. Strokes for Representing Univariate Vector Field Maps. Graphics Interface Proceedings, 249-253. (1989).
- [4] H. Hagh-Shenas, S. Kim, V. Interrante, and C.G. Healey. Weaving Versus Blending: A Quantitative Assessment of the Information Carrying Capacities of Two Alternative Methods for Conveying Multivariate Data With Color." IEEE Transactions on Visualization and Computer Graphics 13, 6, 1270-1277 (2007).
- [5] C.G. Healey, Formalizing Artistic Techniques and Scientific Visualization for Painted Renditions of Complex Information Spaces. In Proceedings International Joint Conference on Artificial Intelligence. 371-376. (2001).
- [6] C.G. Healey, S. Kocherlakota, V. Rao, R. Mehta, and R. St. Amant, "Visual Perception and Mixed-Initiative Interaction for Assisted Visualization Design." IEEE Transactions on Visualization and Computer Graphics, vol. 14, no. 2, pp. 396-411. (2008)
- [7] G. Kindlmann and C.F. Westin. Diffusion Tensor Visualization with Glyph Packing. IEEE Transactions on Visualization and Computer Graphics, 12(5):1329-1336, (2006).
- [8] D. H. Laidlaw, E.T. Ahrens, D. Kramers, M.J. Avalos, C. Readhead, and R.E. Jacobs, Visualizing diffusion tensor images of the mouse spinal cord. Proceedings of the IEEE Visualization Conference. 127-134. (1998).
- [9] M.S. Livingston, and D.H. Hubel, Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. Journal of Neuroscience. 7, 3416-3468. (1987).
- [10] M.S. Livingston, and D.H. Hubel, Segregation of form, movement and depth: Anatomy, physiology and perception. Science 240: 740-749. (1988)
- [11] C. Ware. Color sequences for Univariate Maps: Theory, Experiments and Principles. IEEE Computer Graphics and Applications 8(5), 41-49.(1988).
- [12] C.Ware. Information Visualization: Perception for Design. Morgan Kaufman: San Francisco, (2004).
- [13] C. Ware. Quantitative texton sequences for legible bivariate maps. IEEE Transactions on Visualization and Computer Graphics. 15(6) 1523-1529. (2009).



- [14] T. Watababe and P. Cavanaugh. Texture laciness: the texture equivalent of transparency. *Perception* 25, 293-303, (1966).
- [15] R. Wegenkittl and E. Gröller. Fast Oriented Line Integral Convolution for Vector Field Visualization via the Internet. In *Proceedings IEEE Visualization '97*, pages 309–316, Los Alamitos, Oct. 19–24 (1997).
- [16] Weigle, C., Emigh, W., Liu, G., Taylor, R., Enns, J., and Healey, C. Oriented texture slivers: A technique for local value estimation of multiple scalar fields. *Proc. Graphics Interface 2000*, 163-170. (2000).